1. *What would you choose as the key measure of success of this experiment in encouraging driver partners to serve both cities, and why would you choose this metric?*

As the key measure of success, I would look at the average number of times drivers cross the bridge. This measure seems to be the most straightforward way of determining whether drivers are being exclusive to one city or another. The level to which this number increases or decreases, and whether or not that change is statistically significant, will be the measure of success for the experiment.

2. *Describe a practical experiment you would design to compare the effectiveness of the proposed change in relation to the key measure of success. Please provide details on:*
    a. *how you will implement the experiment*

First I would determine the null (H_0) and alternative (H_1) hypotheses.

H_0 - Covering the cost of the toll has no effect on the number of times drivers cross the bridge.
H_1 - Covering the cost of the toll increases the number of times drivers cross the bridge.

These can be rewritten as:

H_0: $\mu_1 - \mu_2 = 0$
H_1: $\mu_1 - \mu_2 > 0$

*****where μ_1 is the mean of the distribution of bridge crosses when the toll is removed and μ_2 is the mean of the distribution of bridge crosses when the toll is active.

I would then determine a baseline distribution for the number of times drivers cross the bridge while the toll is active. For the purpose of the problem, I will assume the data is normally distributed. I would make sure to use a window of time that controls for anomalous activity, such as holidays, or days when an event that is unique to that date range is taking place. Once I have determined this window, and recorded the baseline distribution, I would identify an analogous window of time, matching the prior window as closely as possible; controlling for factors like seasonality, time-of-month, unusual days, number of drivers working, etc. The experiment would run during this date range, and the distribution of bridge crosses would be recorded. I would then assess whether there is a statistically significant difference between the two distributions.

    b. *what statistical test(s) you will conduct to verify the significance of the observation*

I would use a one-tailed Z-test to determine whether the two population means are significantly different. I chose one-tailed because we expect the proposed change to have a positive impact, and therefore can ignore the possibility of a negative effect, i.e. the second tail. I chose a Z-test in particular because the variances of the populations are known, and the sample size is greater than 30.

c. *how you would interpret the results and provide recommendations to the city operations team along with any caveats.*

I would first determine a significance level; that is, the threshold above which we can confidently reject the null hypothesis. The significance level (or alpha) that I will choose is 5%. In other words, if the mean number of bridge crosses without the toll is high enough that there would be less than a 5% chance of observing it if the null hypothesis were true, then we can conclude that the proposed change affected driver behavior. One caveat to bear in mind is that this does not guarantee that the experiment had an affect; only that there is a reasonably low probability that we are wrong. I would also remind the city operations team that we did our best to control for all external factors, but it would be wise to re-examine this, and try to replicate the results with different date ranges.